

## **RAW SEQUENCE LISTING**

**The Biotechnology Systems Branch of the Scientific and Technical  
Information Center (STIC) no errors detected.**

Application Serial Number: 10/518, 470  
Source: PCT  
Date Processed by STIC: 03/06/2006

# ***ENTERED***



PCT

## RAW SEQUENCE LISTING

DATE: 03/06/2006

PATENT APPLICATION: US/10/518,470

TIME: 16:03:04

Input Set : A:\Sequence Listing\_26473U\_11-11\_05.txt

Output Set: N:\CRF4\03062006\J518470.raw

4 <110> APPLICANT: McGill University  
 6 <120> TITLE OF INVENTION: Oligonucleotide Inhibitors of MBD2/DNA  
 7 Demethylase and Uses Thereof  
 10 <130> FILE REFERENCE: 26473U  
 12 <140> CURRENT APPLICATION NUMBER: 10/518,470  
 C--> 13 <141> CURRENT FILING DATE: 2004-12-20  
 15 <150> PRIOR APPLICATION NUMBER: 60/389,926  
 16 <151> PRIOR FILING DATE: 2002-06-20  
 18 <160> NUMBER OF SEQ ID NOS: 15  
 20 <170> SOFTWARE: FastSEQ for Windows Version 4.0  
 22 <210> SEQ ID NO: 1  
 23 <211> LENGTH: 2584  
 24 <212> TYPE: DNA  
 25 <213> ORGANISM: Homo sapiens  
 27 <220> FEATURE:  
 28 <221> NAME/KEY: misc\_feature  
 29 <222> LOCATION: (0)...(0)  
 30 <223> OTHER INFORMATION: cDNA MBD2/dMTase  
 32 <400> SEQUENCE: 1  
 33 gggggcgctgg ccccgagaag gcggagacaa gatggccgcc catagcgctt ggaggaccta 60  
 34 agaggcggtg gccggggcca cgccccgggc aggaggggcg ctctgtgcgc gcccgcctcta 120  
 35 tgatgcttgc gcgcgtcccc cgcgcgccgc gctgcgggcg gggcggtct cccggattcc 180  
 36 aagggtctcg ttacggaaga agcgcagcgc cggctgggga gggggctgga tgcgcgcgca 240  
 37 cccgggggga ggccgctgct gcccggagca ggaggagggg gagagtgcgc cgggcggcag 300  
 38 cggcgctggc ggcgactccg ccatagagca gggggggccag ggcagcgcgc tcgccccgtc 360  
 39 cccggtgagc ggcgctgcgc gggaaggcgc tcggggcggc ggccgtggcc gggggcggtg 420  
 40 gaagcaggcg ggccggggcg gcggcgtctg tggccgtggc cggggccggg gccgtggccg 480  
 41 gggacgggga cggggccggg gccggggccg cggccgtccc ccgagtggcg gcagcggcct 540  
 42 tggcggcgac ggccggcggt gcggcggcgc cgggcgggcg ggccggcgcg ccccccggcg 600  
 43 ggagccggtc cctttcccg cggggagcgc gggggccggg ccagggggac cccggggccac 660  
 44 ggagagcggg aagaggatgg attgcccgc cttcccccc ggatggaaga aggaggaagt 720  
 45 gatccgaaaa tctgggctaa gtgctggcaa gagcgatgtc tactacttca gtccaagtgg 780  
 46 taagaagttc agaagcaagc ctcaagtggc aaggtacctg ggaaatactg ttgatctcag 840  
 47 cagttttgac ttcagaactg gaaagatgat gcctagtaaa ttacagaaga acaaacagag 900  
 48 actgcgaaac gatcctctca atcaaaataa gggtaaacca gacttgaata caacattgcc 960  
 49 aattagacaa acagcatcaa ttttcaaaca accggtaacc aaagtcacaa atcatcctag 1020  
 50 taataaagtg aaatcagacc cacaacgaat gaatgaacag ccacgtcagc ttttctggga 1080  
 51 gaagaggcta caaggactta gtgcatacaga tgtaacagaa caaattataa aaaccatgga 1140  
 52 actacccaaa ggtcttcaag gagttggtcc aggtagcaat gatgagacct ttttatctgc 1200  
 53 tgttgccagt gctttgcaca caagctctgc gccaatcaca gggcaagtct ccgctgctgt 1260  
 54 ggaaaagaac cctgctgttt ggcttaacac atctcaaccc ctctgcaaag cttttattgt 1320  
 55 cacagatgaa gacatcagga aacaggaaga gcgagtacag caagtacgca agaaattgga 1380  
 56 agaagcactg atggcagaca tctgtcgcg agctgctgat acagaagaga tggatattga 1440

## RAW SEQUENCE LISTING

DATE: 03/06/2006

PATENT APPLICATION: US/10/518,470

TIME: 16:03:04

Input Set : A:\Sequence Listing\_26473U\_11-11\_05.txt

Output Set: N:\CRF4\03062006\J518470.raw

```

57 aatggacagt ggagatgaag cctaagaata tgatcaggta actttcgacc gactttcccc 1500
58 aagagaaaat tcctagaaat tgaacaaaaa tgtttccact ggcttttgcc tgtaagaaaa 1560
59 aaaatgtacc cgagcacata gagcttttta atagcactaa ccaatgcctt tttagatgta 1620
60 tttttgatgt atatatctat tattcaaaaa atcatgttta ttttgagtcc taggacttaa 1680
61 aattagtctt ttgtaatatc aagcaggacc ctaagatgaa gctgagcttt tgatgccagg 1740
62 tgcaatctac tggaaatgta gcacttacgt aaaacatttg tttccccac agttttaata 1800
63 agaacagatc aggaattcta aataaatttc ccagttaaag attattgtga cttcactgta 1860
64 tataaacata tttttatact ttattgaaag gggacacctg tacattcttc catcatcact 1920
65 gtaaagacaa ataaatgatt atattcacag actgattgga attctttctg ttgaaaagca 1980
66 cacacaataa agaacccttc gttagccttc ctctgattta cattcaactc tgatccctgg 2040
67 gccttaggtt tgacatggag gtggaggaag atagcgcata tatttgagc atgaactatt 2100
68 gcctctggac gttgtgagaa ttgtgctttc accagaattt ctaagaattt ctgctaaata 2160
69 tcacctagca tgtgtaattt tttttccttg cctgtgactt ggacttttga tagttctata 2220
70 agaataaggc tttttcttcc cttgggcatg agtcagatac acaaggaccc ttcagggtgt 2280
71 actagaaggc gtccatgttt attgtttttt aaagaatgtt tggcactctc taacgtccac 2340
72 tagcttactg agttatcagg tgcaggtcag actcttggtt acagtgaag gcagcttcta 2400
73 ggcagagttg cttaatgaaa gggtttgtaa tactttacaa accattacct gtacctggcc 2460
74 tggcctccaa aatattaaca ttctttttct gttgaaactc gcgagtgtaa ctttcatacc 2520
75 acttgaattt attgatattt aattatgaaa actagcatta cattattaaa cgatttctaa 2580
76 aatc

```

78 &lt;210&gt; SEQ ID NO: 2

79 &lt;211&gt; LENGTH: 411

80 &lt;212&gt; TYPE: PRT

81 &lt;213&gt; ORGANISM: Homo sapiens

83 &lt;400&gt; SEQUENCE: 2

```

84 Met Arg Ala His Pro Gly Gly Gly Arg Cys Cys Pro Glu Gln Glu Glu
85 1 5 10 15
86 Gly Glu Ser Ala Ala Gly Gly Ser Gly Ala Gly Gly Asp Ser Ala Ile
87 20 25 30
88 Glu Gln Gly Gly Gln Gly Ser Ala Leu Ala Pro Ser Pro Val Ser Gly
89 35 40 45
90 Val Arg Arg Glu Gly Ala Arg Gly Gly Gly Arg Gly Arg Gly Arg Trp
91 50 55 60
92 Lys Gln Ala Gly Arg Gly Gly Gly Val Cys Gly Arg Gly Arg Gly Arg
93 65 70 75 80
94 Gly Arg Gly Arg Gly Arg Gly Arg Gly Arg Gly Arg Gly Arg Gly Arg
95 85 90 95
96 Pro Pro Ser Gly Gly Ser Gly Leu Gly Gly Asp Gly Gly Gly Cys Gly
97 100 105 110
98 Gly Gly Gly Ser Gly Gly Gly Gly Ala Pro Arg Arg Glu Pro Val Pro
99 115 120 125
100 Phe Pro Ser Gly Ser Ala Gly Pro Gly Pro Arg Gly Pro Arg Ala Thr
101 130 135 140
102 Glu Ser Gly Lys Arg Met Asp Cys Pro Ala Leu Pro Pro Gly Trp Lys
103 145 150 155 160
104 Lys Glu Glu Val Ile Arg Lys Ser Gly Leu Ser Ala Gly Lys Ser Asp
105 165 170 175
106 Val Tyr Tyr Phe Ser Pro Ser Gly Lys Lys Phe Arg Ser Lys Pro Gln
107 180 185 190

```

## RAW SEQUENCE LISTING

DATE: 03/06/2006

PATENT APPLICATION: US/10/518,470

TIME: 16:03:04

Input Set : A:\Sequence Listing\_26473U\_11-11\_05.txt

Output Set: N:\CRF4\03062006\J518470.raw

```

108 Leu Ala Arg Tyr Leu Gly Asn Thr Val Asp Leu Ser Ser Phe Asp Phe
109      195      200      205
110 Arg Thr Gly Lys Met Met Pro Ser Lys Leu Gln Lys Asn Lys Gln Arg
111      210      215      220
112 Leu Arg Asn Asp Pro Leu Asn Gln Asn Lys Gly Lys Pro Asp Leu Asn
113 225      230      235      240
114 Thr Thr Leu Pro Ile Arg Gln Thr Ala Ser Ile Phe Lys Gln Pro Val
115      245      250      255
116 Thr Lys Val Thr Asn His Pro Ser Asn Lys Val Lys Ser Asp Pro Gln
117      260      265      270
118 Arg Met Asn Glu Gln Pro Arg Gln Leu Phe Trp Glu Lys Arg Leu Gln
119      275      280      285
120 Gly Leu Ser Ala Ser Asp Val Thr Glu Gln Ile Ile Lys Thr Met Glu
121      290      295      300
122 Leu Pro Lys Gly Leu Gln Gly Val Gly Pro Gly Ser Asn Asp Glu Thr
123 305      310      315      320
124 Leu Leu Ser Ala Val Ala Ser Ala Leu His Thr Ser Ser Ala Pro Ile
125      325      330      335
126 Thr Gly Gln Val Ser Ala Ala Val Glu Lys Asn Pro Ala Val Trp Leu
127      340      345      350
128 Asn Thr Ser Gln Pro Leu Cys Lys Ala Phe Ile Val Thr Asp Glu Asp
129      355      360      365
130 Ile Arg Lys Gln Glu Glu Arg Val Gln Gln Val Arg Lys Lys Leu Glu
131      370      375      380
132 Glu Ala Leu Met Ala Asp Ile Leu Ser Arg Ala Ala Asp Thr Glu Glu
133 385      390      395      400
134 Met Asp Ile Glu Met Asp Ser Gly Asp Glu Ala
135      405      410

```

138 &lt;210&gt; SEQ ID NO: 3

139 &lt;211&gt; LENGTH: 1953

140 &lt;212&gt; TYPE: DNA

141 &lt;213&gt; ORGANISM: Mus musculus

143 &lt;220&gt; FEATURE:

144 &lt;221&gt; NAME/KEY: misc\_feature

145 &lt;222&gt; LOCATION: (0)...(0)

146 &lt;223&gt; OTHER INFORMATION: cDNA MBD2/dMTase

148 &lt;400&gt; SEQUENCE: 3

```

149 ggggggctg gcccagagga ggcggagaca atatggcctc gccctagct tggaggacct 60
150 aagaggcgcg gccggggcca cgccccgggc gggagggccg ctctgtgcgc gcccgctcta 120
151 tgatgcttgc gcgcgtcccc cgcgcgccgc tctgcgggcg gggcggtct ccgggattcc 180
152 aagggtcgg ttacggaaga agcgagagc cggctgggga gggggtgga tgcgcgcgca 240
153 cccgggggga ggccgtgct gcccgagca ggaggaggg gagagcgcg cggcggcgag 300
154 cggcgtggc ggcgactcc ccatagagca ggggggccag ggcagcgcg tcgctccgtc 360
155 cccggtgagc ggcgtgcga ggaaggcgc tcggggcggc gccgtggcc gggggcggtg 420
156 gaagcaggcg gccggggcg gcggcgctc tggccgtggc cgtggccgtg gccggggtcg 480
157 gggccgtggc cggggccggg gccggggccg cggccgtccc cagagtggcg gcagcggcct 540
158 tggcggcgac ggcgcgcgcg gcgcggcgcg ctgcggcgct gccagcggtg gcggcgctcg 600
159 cccccggcgg gatcctgtcc ctttcccgtc gggagctcg gggccggggc ccaggggacc 660
160 ccgggccacg gagagcggga agaggatgga ctgcccgcc ccccccccg gatggaagaa 720

```

## RAW SEQUENCE LISTING

DATE: 03/06/2006

PATENT APPLICATION: US/10/518,470

TIME: 16:03:04

Input Set : A:\Sequence Listing\_26473U\_11-11\_05.txt

Output Set: N:\CRF4\03062006\J518470.raw

```

161 ggaggaagtg atccgaaaat cagggtcag tgctggcaag agcgatgtct actacttcag 780
162 tccaagtggg aagaagttca gaagtaaacc tcagctggca agatacctgg gaaatgctgt 840
163 tgaccttagc agttttgact tcaggaccgg caagatgatg cctagtaaata tacagaagaa 900
164 caagcagaga ctccggaatg acccctcaa tcagaacaag ggtaaaccag acctgaacac 960
165 aacattgcca attagacaaa ctgcatcaat tttcaagcaa ccagtaacca aattcacgaa 1020
166 ccacccgagc aataaggtga agtcagaccc ccagcggatg aatgaacaac cacgtcagct 1080
167 tttctgggag aagaggctac aaggacttag cgcacagat gtaacagaac aaattataaa 1140
168 aacctgggag ctacctaaag gtcttcaagg agtcgggtcca ggtagcaatg acgagacct 1200
169 tctgtctgct gtggccagtg ctttacacac aagctctgcg cccatcacag gacaagtctc 1260
170 tgctgccgtg gaaaagaacc ctgctgtttg gcttaacaca tctcaacccc tctgcaaagc 1320
171 tttcattggt acagatgaag acattaggaa acaggaagag cgagtccaac aagtacgcaa 1380
172 gaaactggag gaggcactga tggccgacat cctgtcccgg gctgcggaca cggaggaagt 1440
173 agacattgac atggacagtg gagatgaggc gtaagaatat gatcaggtaa ctttcgactg 1500
174 accttcccc aagacaaatt gctagaaaca gaattaaaac atttccactg ggtttcgcct 1560
175 gtaagaaaaa gtgtacctga gcacatagct ttttaatagc actaaccaat gcctttttag 1620
176 atgtattttt gatgtatata tctattattc caaatgatgt ttattttgaa tcctaggact 1680
177 taaaatgagt cttttataat agcaagcagg gcccttcggg tgcagtgcag ctttgaggcc 1740
178 aggtgcagtc tactggaaag gtagcactta cgtgaaatat ttgtttcccc cacagtttta 1800
179 atataaacag atcaggagta ccaaataagt ttcccaatta aagattatta tacttctactg 1860
180 tatataaaca gatttttata ctttattgaa agaagatacc tgtacattct tccatcatca 1920
181 ctgtaaagac aaataaatga ctatattcac aga 1953

```

183 &lt;210&gt; SEQ ID NO: 4

184 &lt;211&gt; LENGTH: 414

185 &lt;212&gt; TYPE: PRT

186 &lt;213&gt; ORGANISM: Mus musculus

188 &lt;400&gt; SEQUENCE: 4

```

189 Met Arg Ala His Pro Gly Gly Gly Arg Cys Cys Pro Glu Gln Glu Glu
190 1 5 10 15
191 Gly Glu Ser Ala Ala Gly Gly Ser Gly Ala Gly Gly Asp Ser Ala Ile
192 20 25 30
193 Glu Gln Gly Gly Gln Gly Ser Ala Leu Ala Pro Ser Pro Val Ser Gly
194 35 40 45
195 Val Arg Arg Glu Gly Ala Arg Gly Gly Gly Arg Gly Arg Gly Arg Trp
196 50 55 60
197 Lys Gln Ala Ala Arg Gly Gly Gly Val Cys Gly Arg Gly Arg Gly Arg
198 65 70 75 80
199 Gly Arg Gly Arg Gly Arg Gly Arg Gly Arg Gly Arg Gly Arg Gly Arg
200 85 90 95
201 Pro Gln Ser Gly Gly Ser Gly Leu Gly Gly Asp Gly Gly Gly Gly Ala
202 100 105 110
203 Gly Gly Cys Gly Val Gly Ser Gly Gly Gly Val Ala Pro Arg Arg Asp
204 115 120 125
205 Pro Val Pro Phe Pro Ser Gly Ser Ser Gly Pro Gly Pro Arg Gly Pro
206 130 135 140
207 Arg Ala Thr Glu Ser Gly Lys Arg Met Asp Cys Pro Ala Leu Pro Pro
208 145 150 155 160
209 Gly Trp Lys Lys Glu Glu Val Ile Arg Lys Ser Gly Leu Ser Ala Gly
210 165 170 175
211 Lys Ser Asp Val Tyr Tyr Phe Ser Pro Ser Gly Lys Lys Phe Arg Ser

```

## RAW SEQUENCE LISTING

DATE: 03/06/2006

PATENT APPLICATION: US/10/518,470

TIME: 16:03:04

Input Set : A:\Sequence Listing\_26473U\_11-11\_05.txt

Output Set: N:\CRF4\03062006\J518470.raw

```

212          180          185          190
213 Lys Pro Gln Leu Ala Arg Tyr Leu Gly Asn Ala Val Asp Leu Ser Ser
214          195          200          205
215 Phe Asp Phe Arg Thr Gly Lys Met Met Pro Ser Lys Leu Gln Lys Asn
216          210          215          220
217 Lys Gln Arg Leu Arg Asn Asp Pro Leu Asn Gln Asn Lys Gly Lys Pro
218 225          230          235          240
219 Asp Leu Asn Thr Thr Leu Pro Ile Arg Gln Thr Ala Ser Ile Phe Lys
220          245          250          255
221 Gln Pro Val Thr Lys Phe Thr Asn His Pro Ser Asn Lys Val Lys Ser
222          260          265          270
223 Asp Pro Gln Arg Met Asn Glu Gln Pro Arg Gln Leu Phe Trp Glu Lys
224          275          280          285
225 Arg Leu Gln Gly Leu Ser Ala Ser Asp Val Thr Glu Gln Ile Ile Lys
226          290          295          300
227 Thr Met Glu Leu Pro Lys Gly Leu Gln Gly Val Gly Pro Gly Ser Asn
228 305          310          315          320
229 Asp Glu Thr Leu Leu Ser Ala Val Ala Ser Ala Leu His Thr Ser Ser
230          325          330          335
231 Ala Pro Ile Thr Gly Gln Val Ser Ala Ala Val Glu Lys Asn Pro Ala
232          340          345          350
233 Val Trp Leu Asn Thr Ser Gln Pro Leu Cys Lys Ala Phe Ile Val Thr
234          355          360          365
235 Asp Glu Asp Ile Arg Lys Gln Glu Glu Arg Val Gln Gln Val Arg Lys
236          370          375          380
237 Lys Leu Glu Glu Ala Leu Met Ala Asp Ile Leu Ser Arg Ala Ala Asp
238 385          390          395          400
239 Thr Glu Glu Val Asp Ile Asp Met Asp Ser Gly Asp Glu Ala
240          405          410

```

243 &lt;210&gt; SEQ ID NO: 5

244 &lt;211&gt; LENGTH: 18

245 &lt;212&gt; TYPE: DNA

246 &lt;213&gt; ORGANISM: Artificial Sequence

248 &lt;220&gt; FEATURE:

249 &lt;223&gt; OTHER INFORMATION: Antisense oligonucleotide

251 &lt;400&gt; SEQUENCE: 5

252 ggcaatccat cctcttcc

18

254 &lt;210&gt; SEQ ID NO: 6

255 &lt;211&gt; LENGTH: 18

256 &lt;212&gt; TYPE: DNA

257 &lt;213&gt; ORGANISM: Artificial Sequence

259 &lt;220&gt; FEATURE:

260 &lt;223&gt; OTHER INFORMATION: Antisense oligonucleotide

262 &lt;400&gt; SEQUENCE: 6

263 cttcctcctt cttccatc

18

265 &lt;210&gt; SEQ ID NO: 7

266 &lt;211&gt; LENGTH: 17

267 &lt;212&gt; TYPE: DNA

268 &lt;213&gt; ORGANISM: Artificial Sequence

**VERIFICATION SUMMARY**

PATENT APPLICATION: US/10/518,470

DATE: 03/06/2006

TIME: 16:03:05

Input Set : A:\Sequence Listing\_26473U\_11-11\_05.txt

Output Set: N:\CRF4\03062006\J518470.raw

L:13 M:271 C: Current Filing Date differs, Replaced Current Filing Date